

日本国特許庁
JAPAN PATENT OFFICE

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office.

出願年月日 2002年10月31日
Date of Application:

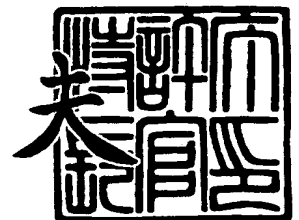
出願番号 特願2002-318627
Application Number:
[ST. 10/C]: [JP 2002-318627]

出願人 セイコーエプソン株式会社
Applicant(s):

2003年 8月12日

特許庁長官
Commissioner,
Japan Patent Office

今井 康



【書類名】 特許願

【整理番号】 J0093425

【あて先】 特許庁長官殿

【国際特許分類】 G10L 3/00

【発明者】

【住所又は居所】 長野県諏訪市大和3丁目3番5号 セイコーエプソン株式会社内

【氏名】 宮澤 康永

【特許出願人】

【識別番号】 000002369

【氏名又は名称】 セイコーエプソン株式会社

【代理人】

【識別番号】 100095728

【弁理士】

【氏名又は名称】 上柳 雅誉

【連絡先】 0 2 6 6 - 5 2 - 3 1 3 9

【選任した代理人】

【識別番号】 100107076

【弁理士】

【氏名又は名称】 藤網 英吉

【選任した代理人】

【識別番号】 100107261

【弁理士】

【氏名又は名称】 須澤 修

【手数料の表示】

【予納台帳番号】 013044

【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【包括委任状番号】 0109826

【プルーフの要否】 要

【書類名】 明細書

【発明の名称】 音声認識用の音響モデル作成方法および音声認識装置

【特許請求の範囲】

【請求項 1】 運転者の操作によって走行可能な車両内で音声認識を行う音声認識用の音響モデル作成方法であって、

前記車両をある期間走行させることによって、その車両内で収集可能な各種雑音を時系列で収集し、その収集された雑音をある統計的手法によってクラスタリングして n 種類 (n は正の整数) の雑音データを作成し、その n 種類の雑音データを予め用意された標準的な音声データに重畳させて n 種類の雑音重畳音声データを作成し、その n 種類の雑音重畳音声データに対して雑音除去処理を行って n 種類の雑音除去音声データを作成し、その n 種類の雑音除去音声データから n 種類の音響モデルを作成することを特徴とする音声認識用の音響モデル作成方法。

【請求項 2】 前記 n 種類の雑音重畳音声データに対して行われる雑音除去処理は、個々の雑音データに適した雑音除去手法を用いて行うことを特徴とする請求項 1 記載の音声認識用の音響モデル作成方法。

【請求項 3】 前記車両内で収集可能な各種雑音は、車両の走行状況、走行位置、天候、車両に搭載されている機器の動作状況に起因する音であることを特徴とする請求項 1 または 2 記載の音声認識用の音響モデル作成方法。

【請求項 4】 前記車両内で収集可能な各種雑音を時系列で収集する際、車両の走行状況、走行位置、天候、車両に搭載されている機器の動作状況を表す情報を時系列で記録することを特徴とする請求項 1 から 3 のいずれかに記載の音声認識用の音響モデル作成方法。

【請求項 5】 運転者の操作によって走行可能な車両内で音声認識を行う音声認識装置であって、

認識すべき音声およびそれ以外の雑音の入力が可能な音入力手段と、

当該車両の走行状況、走行位置、天候、車両に搭載されている機器の動作状況を表す情報を入力し、入力された情報に基づいて前記音入力手段から入力された現在の雑音が、予めクラスタリングされた n 種類 (n は正の整数) の雑音データのうちのどの雑音データに属する雑音であるかを判別する雑音種類判別手段と、

前記 n 種類の雑音データをそれぞれ標準的な音声データに重畳させて n 種類の雑音重畳音声データを作成し、その n 種類の雑音重畳音声データからそれぞれ雑音除去処理を行って n 種類の雑音除去音声データを作成して、その n 種類の雑音除去音声データに基づいて作成された n 種類の音響モデルと、

前記音入力手段から雑音の重畳された雑音重畳音声データを入力するとともに、前記雑音種類判別手段からの雑音種類判別結果を入力し、雑音重畳音声データに対して前記雑音種類判別手段での判別結果に基づいて雑音除去を行う雑音除去処理手段と、

この雑音除去処理手段によって雑音除去された雑音除去音声に対し、前記 n 種類の音響モデルのうち前記雑音種類判別手段で判別された雑音種類に対応した音響モデルを用いて音声認識を行う音声認識手段と、

を有したことを特徴とする音声認識装置。

【請求項 6】 前記クラスタリングによって得られた n 種類の雑音データに対して行われる雑音除去処理は、個々の雑音データに適した雑音除去手法を用いて行うことを特徴とする請求項 5 記載の音声認識装置。

【請求項 7】 前記 n 種類の雑音データは、前記車両をある期間走行させることによってその車両内で収集可能な各種音を時系列で収集し、その収集された音のある統計的手法によってクラスタリングすることによって得られることを特徴とする請求項 5 または 6 記載の音声認識装置。

【請求項 8】 前記車両内で収集可能な各種雑音は、車両の走行状況、走行位置、天候、車両に搭載されている機器の動作状況に起因する音であることを特徴とする請求項 5 から 7 のいずれかに記載の音声認識装置。

【請求項 9】 前記車両内で収集可能な各種雑音を時系列で収集する際、車両の走行状況、走行位置、天候、車両に搭載されている機器の動作状況を表す情報を時系列で記録することを特徴とする請求項 5 から 8 のいずれかに記載の音声認識装置。

【請求項 10】 前記 n 種類の音響モデルを作成する際の雑音除去処理と前記認識すべき音声に対する音声認識を行う際の雑音除去処理は同じ雑音除去手法を用いることを特徴とする請求項 5 から 9 のいずれかに記載の音声認識装置。

【発明の詳細な説明】**【0001】****【発明の属する技術分野】**

本発明は、たとえば自動車などの車内で音声認識を行うための音声認識用の音響モデル作成方法および音声認識装置に関し、特に、カーナビゲーション、カーオーディオ、カーエアコンディショナなどの車載機器を音声によって操作する際の音声認識を行う音声認識用の音響モデル作成方法および音声認識装置に関する。

【0002】**【従来の技術】**

最近、音声認識技術は様々な分野で利用され、各種機器を音声によって操作可能とすることが一般的に行われるようになってきている。このように、音声によってある特定の機器の操作を可能とすることで、両手で何らかの操作を行っているときに、他の機器操作を行う必要のある場合などにおいてきわめて便利なものとなる。

【0003】

たとえば、自動車に搭載されているカーナビゲーション、カーオーディオ、カーエアコンディショナ（以下ではエアコンという）などの様々な機器は、通常は、必要なときに運転手が自らの手を使って操作するのが一般的であるが、最近では、これらの機器を音声で操作可能とする技術が種々提案され実用化されつつある。これによって、運転中であっても、ステアリングから手を離すことなく、これらの機器のスイッチをオン・オフさせたり、それらの機器の機能設定が可能となるので、安全運転にもつながり、今後、ますます普及して行くものと考えられる。

【0004】

しかし、自動車などに搭載される上述のような機器の操作を音声によって行うには、様々な雑音の存在する環境下で高い認識性能が得られるようにすることが重要であり、この点が従来からの大きな課題ともなっている。

【0005】

このように、自動車の車内など様々な雑音の存在する環境下で音声認識を行う方法として、従来では、図8に示すような方法で音響モデルを作成し、その音響モデルを用いて図9に示すように音声認識を行っている。

【0006】

この従来の音響モデル作成処理について図8を参照しながら説明する。まず、無響室などの雑音のない環境で収集された標準的な音声データ（たとえば、多数の話者が多種類単語について発話して得られた大量の音声データ） V と、ある特定種類の雑音データ N を雑音重畳データ作成部51に入力して、その標準的な音声データにある特定種類の雑音をある S/N 比で重畳させた雑音重畳音声データ VN を作成する。

【0007】

この雑音重畳音声データ VN に対し、雑音除去処理部52がスペクトラム・サブトラクション（ SS ）やケプストラム平均化処理（ CMN ）など、その雑音の種類に最適な雑音除去処理を施し、雑音除去音声データ V' （雑音除去処理を行っても除去されない雑音成分が残る）を作成する。そして、音響モデル学習処理部53がこの雑音除去音声データ V' を用いて音素HMM（Hidden Markov Model）や音節HMMなどの音響モデル M を作成する。

【0008】

一方、音声認識処理は図9に示すように、マイクロホン61から入力された話者の音声データ（機器操作の音声コマンド）に対し、入力信号処理部62が増幅や A/D 変換（アナログ／デジタル変換）などを行い、そのあと、雑音除去処理部63が入力音声データに対して雑音除去処理（図8の雑音除去処理部52で行ったと同じ手法による雑音除去処理）を行う。

【0009】

そして、その雑音除去された音声データ（雑音除去音声データという）に対して、音声認識処理部64が言語モデル65と図8の音響モデル学習処理部53で作成された音響モデル M を用いて音声認識処理する。

【0010】

【特許文献1】

特開 2002-132289 号公報

【0011】

【発明が解決しようとする課題】

以上が従来から一般的に行われているある特定の雑音環境下での音声認識手法である。このように、自動車の車内で音声認識を行う場合、自動車の走行状況に起因する音（速度などに応じたタイヤのパターンノイズや窓の開度に応じた風切り音、回転数や変速ギヤ位置などによるエンジン音）、周囲の環境に起因する音（トンネルなどを通過する際の反響音など）、自動車に搭載されている機器の動作状況に起因する音（カーオーディオの動作音やエアコンディショナの動作音、ワイパーや方向指示器の動作音など）、降雨時における雨音など自動車特有の雑音がマイクロホン 61 から入り、これらの雑音が音声コマンドに重畳した状態で以降の処理部に渡されることになる。

【0012】

なお、自動車の場合、マイクロホンから入力される雑音の種類は上述したように、自動車特有の雑音でありその種類はある程度限定されるが、走行状況に起因する雑音としてエンジン音だけに注目して考えても、エンジン始動前の停車時やエンジン始動後であってもアイドリング時、低速走行時、高速走行時では雑音の大きさや雑音の種類が異なる場合が多い。また、同じ速度で走行していても、変速ギヤ比の関係でエンジンの回転数が高いときと低いときで雑音の大きさや種類が異なるなど、エンジン音だけを考えても、雑音の大きさと種類は様々異なってくる。さらに、このような走行状況に起因するものだけではなく、上述したように、たとえば、窓の開閉度合いによる風切り音、トンネルや橋など周囲の建造物などの反響音、降雨時（降雨量によっても異なる）による雨音、エアコン、ワイパー、カーオーディオ、方向指示器など各種機器の動作音も雑音としてマイクロホンに入力されてくる。

【0013】

しかし、上述した従来の音声認識手法では、ある特定の雑音のみに対応して作成された音響モデルのみを使って音声認識するようにしているので、上述したような刻々と変化する多種多様な雑音には対応できず、状況によって発生する雑音

が音声認識性能に大きく影響を与え、高い認識率を得ることは難しいという問題がある。

【0014】

そこで本発明は、たとえば、自動車などに搭載された機器の操作を音声によって行う際、刻々と変化する走行状況、走行位置、天候、搭載機器の動作状況などに起因する雑音に対応した音声認識を可能とし、多種多様な雑音環境下で高い認識性能を可能とすることを目的としている。

【0015】

【課題を解決するための手段】

上述した目的を達成するために本発明の音声認識用の音響モデル作成方法は、運転者の操作によって走行可能な車両内で音声認識を行う音声認識用の音響モデル作成方法であって、前記車両をある期間走行させることによって、その車両内で収集可能な各種雑音を時系列で収集し、その収集された雑音をある統計的手法によってクラスタリングして n 種類の雑音データを作成し、その n 種類の雑音データを予め用意された標準的な音声データに重畳させて n 種類の雑音重畳音声データを作成し、その n 種類の雑音重畳音声データに対して雑音除去処理を行って n 種類の雑音除去音声データを作成し、その n 種類の雑音除去音声データから n 種類の音響モデルを作成するようにしている。

【0016】

このような音声認識用の音響モデル作成方法において、前記 n 種類の雑音重畳音声データに対して行われる雑音除去処理は、個々の雑音データに適した雑音除去手法を用いて行うようにしている。

【0017】

また、前記車両内で収集可能な各種雑音は、車両の走行状況、走行位置、天候、車両に搭載されている機器の動作状況に起因する音である。

【0018】

また、前記車両内で収集可能な各種雑音を時系列で収集する際、車両の走行状況、走行位置、天候、車両に搭載されている機器の動作状況を表す情報を時系列で記録するようにしている。

【0019】

また、本発明の音声認識装置は、運転者の操作によって走行可能な車両内で音声認識を行う音声認識装置であって、認識すべき音声およびそれ以外の雑音の入力が可能な音入力手段と、当該車両の走行状況、走行位置、天候、車両に搭載されている機器の動作状況を表す情報を入力し、入力された情報に基づいて前記音入力手段から入力された現在の雑音が、予めクラスタリングされた n 種類（ n は正の整数）の雑音データのうちの雑音データに属する雑音であるかを判別する雑音種類判別手段と、前記 n 種類の雑音データをそれぞれ標準的な音声データに重畳させて n 種類の雑音重畳音声データを作成し、その n 種類の雑音重畳音声データからそれぞれ雑音除去処理を行って n 種類の雑音除去音声データを作成して、その n 種類の雑音除去音声データに基づいて作成された n 種類の音響モデルと、前記音入力手段から雑音の重畳された雑音重畳音声データを入力するとともに、前記雑音種類判別手段からの雑音種類判別結果を入力し、雑音重畳音声データに対して前記雑音種類判別手段での判別結果に基づいて雑音除去を行う雑音除去処理手段と、この雑音除去処理手段によって雑音除去された雑音除去音声に対し、前記 n 種類の音響モデルのうち前記雑音種類判別手段で判別された雑音種類に対応した音響モデルを用いて音声認識を行う音声認識手段とを有している。

【0020】

このような音声認識装置において、前記クラスタリングによって得られた n 種類の雑音データに対して行われる雑音除去処理は、個々の雑音データに適した雑音除去手法を用いて行う。

【0021】

また、前記 n 種類の雑音データは、前記車両をある期間走行させることによってその車両内で収集可能な各種音を時系列で収集し、その収集された音をある統計的手法によってクラスタリングすることによって得られる。

【0022】

また、前記車両内で収集可能な各種雑音は、車両の走行状況、走行位置、天候、車両に搭載されている機器の動作状況に起因する音である。

【0023】

また、前記車両内で収集可能な各種雑音を時系列で収集する際、車両の走行状況、走行位置、天候、車両に搭載されている機器の動作状況を表す情報を時系列で記録するようにしている。

【0024】

また、前記 n 種類の音響モデルを作成する際の雑音除去処理と前記認識すべき音声に対する音声認識を行う際の雑音除去処理は同じ雑音除去手法を用いる。

【0025】

このように本発明によれば、自動車などの車両をある期間走行させることによって、その車両内で収集可能な各種雑音を時系列で収集し、その収集された雑音のある統計的手法によってクラスタリングして n 種類の雑音データを作成し、その n 種類の雑音データを予め用意された標準的な音声データに重畳させて n 種類の雑音重畳音声データを作成し、その n 種類の雑音重畳音声データに対して雑音除去処理を行って n 種類の雑音除去音声データを作成し、その n 種類の雑音除去音声データから n 種類の音響モデルを作成するようにしているので、車両の走行状況、走行位置、天候、車両内の機器の動作状況などに起因する様々な雑音の種類に対応した最適な音響モデルを作成することができる。このような雑音の種類に対応した最適な音響モデルを用いて音声認識を行うことにより、様々な雑音環境下で高い認識性能を得ることができる。

【0026】

また、前記 n 種類の雑音重畳音声データに対して行われる雑音除去処理は、個々の雑音データに適した雑音除去手法を用いることによって、それぞれの雑音に対して適正かつ効率的な雑音除去が可能となる。

【0027】

また、前記車両内で収集可能な各種雑音は、車両の走行状況、走行位置、天候、車両に搭載されている機器の動作状況に起因する音であって具体的には、走行速度、エンジン回転数、変速ギヤ位置によるエンジン音やタイヤのパターンノイズ、トンネル内における反響音、降雨時における雨音、ワイパー、方向指示器、カーオーディオ機器の動作音であり、これらの音が雑音として収集され、これらの雑音をクラスタリングして、それぞれの雑音クラスタに対応した雑音データを

生成し、それぞれの雑音データごとの音響モデルを作成することによって、車両特有の様々な雑音に対応可能な音響モデルを作成することができる。

【0028】

また、前記車両内で収集可能な各種雑音を時系列で収集する際、車両の走行状況、走行位置、天候、車両に搭載されている機器の動作状況を表す情報を時系列で記録するようにしている。このように、たとえば、走行速度、エンジン回転数、変速ギヤ位置、窓の開閉状態、走行場所、カーエアコンやカーオーディオの動作状況などをそれぞれ時系列で記録することによって、収集される雑音とその雑音の直接的または間接的な発生源との対応付けが可能となる。

【0029】

また、本発明の音声認識装置は、上述した音響モデル作成方法によって作成されたn種類の音響モデルを持っている。そして、現在の雑音が予めクラスタリングされたn種類の雑音データのうち、どの雑音データに属する雑音であるかを判別する雑音種類判別を行い、また、音声入力手段から雑音の重畳された雑音重畳音声データを入力するとともに、前記雑音種類判別結果を入力し、雑音重畳音声に対してその雑音種類判別結果に基づいた雑音除去を行い、その雑音除去音声データに対し、前記n種類の音響モデルのうち前記雑音種類判別部で判別された雑音種類に対応した音響モデルを用いて音声認識を行うようにしているので、自動車など車両の走行状況、走行位置、天候、車両に搭載されている機器の動作状況に起因する雑音に対し、最適な雑音除去処理が行えとともに、最適な音響モデルを用いた音声認識が可能となる。これによって、自動車の車内などのように、ある程度限られた種類の雑音が存在する環境下で音声認識する場合、雑音の種類に適合した音響モデルを用いた音声認識が可能となるので、高い認識性能を得ることができる。

【0030】

【発明の実施の形態】

以下、本発明の実施の形態について説明する。なお、この実施の形態で説明する内容は、本発明の音声認識用の音響モデル作成方法および音声認識装置の両方の説明を含むものである。また、この実施の形態では、自動車に搭載された各種

機器の操作を音声によって行う場合を例にとって説明する。まず、音声認識用の音響モデル作成方法の大まかな処理手順について図1を参照しながら簡単に説明する。

【0031】

自動車の場合、音声コマンド入力用のマイクロホンに入力される雑音の種類は自動車特有のものが多く、しかも、その雑音をあらかじめ収集しておくこともできる。

【0032】

そこで、自動車の室内で音声認識を行う際、音声認識性能に影響を与えそうな自動車特有の多種多様な雑音を収集し、収集された多種多様な雑音を統計的手法によってクラスタリングして、 n 個の雑音クラスタを作成し、それぞれの雑音クラスタごとの n 種類の雑音データ N_1, N_2, \dots, N_n を作成する（これについて後に詳細に説明する）。なお、この n 種類の雑音データ N_1, N_2, \dots, N_n には、 S/N 比の違いも考慮されている。たとえば、同じ雑音でも S/N 比が0デシベルから20デシベル程度に広がっている場合、 S/N 比の違いで雑音をクラスタリングして、 n 個の雑音クラスタに分けて、その n 個の雑音クラスタごとの雑音データ（ n 種類の雑音データ N_1, N_2, \dots, N_n ）を作成する。

【0033】

そして、無響室などで収集された標準的な音声データ V （たとえば、多数の話者が多種類の単語について発話して得られた大量の音声データ）と上述の n 種類の雑音データ N_1, N_2, \dots, N_n とを雑音重畳音声データ作成部1に与え、標準的な音声データ V と上述の n 種類の雑音データ N_1, N_2, \dots, N_n とをそれぞれ重畳させて n 種類の雑音重畳音声データ VN_1, VN_2, \dots, VN_n を作成する。

【0034】

そして、雑音除去処理部2が n 種類の雑音重畳音声データ VN_1, VN_2, \dots, VN_n に対して最適な雑音除去処理手法を用いて雑音除去処理を行い、 n 種類の雑音除去音声データ V_1', V_2', \dots, V_n' を作成する。その後

、音響モデル学習処理部 3 がこの n 種類の雑音除去音声データ V_1' , V_2' , \dots , V_n' を用いて音響モデルの学習を行い、 n 種類の音響モデル M_1 , M_2 , \dots , M_n を作成する。

【0035】

なお、 n 種類の雑音重畳音声データ V_{N1} , V_{N2} , \dots , V_{Nn} に対するそれぞれ最適な雑音除去処理手法としては、 n 種類の雑音重畳音声データ V_{N1} , V_{N2} , \dots , V_{Nn} ごとに用意された n 種類の雑音除去処理を用いて行うようにしてもよいが、数種類の代表的な雑音除去処理手法を用意して、その中からそれぞれの雑音重畳音声データにとって最適と思われる雑音除去処理手法を選択して用いるようにしてもよい。

【0036】

数種類の代表的な雑音除去処理手法としては、たとえば、前述したようなスペクトラム・サブストラクション (SS) やケプストラム平均化処理 (CMN)、さらには、音源を推定したエコーキャンセルなどがあり、これら雑音除去処理手法のうち、それぞれの雑音に最適な1つの雑音除去手法を選んで雑音除去を行うようにしてもよく、また、これらの雑音除去手法のうち2種類以上を組み合わせ、組み合わせたそれぞれの雑音除去手法に重み付けを行って雑音除去するようにしてもよい。

【0037】

次に、収集された多種多様な雑音を、ある統計的手法によって幾つか (n 個) にクラスタリングし、それぞれの雑音クラスごとの n 種類の雑音データ N_1 , N_2 , \dots , N_n を生成する具体的な例について図 2 を参照しながら詳細に説明する。

【0038】

この実施の形態では、自動車に搭載された機器を操作するための音声コマンドを認識する際に本発明を適用する例であるので、雑音データ収集用の自動車を様々な条件で様々な状況のもとに長期間走行させて、自動車内の所定の場所に設置されたマイクロホン 11 から自動車特有の多種多様な雑音データを時系列で収集する。

【0039】

なお、このマイクロホン11は、雑音データ収集用の自動車内において、運転者が音声によって機器操作を行う際、話者の音声コマンドを適正に入力できるような位置に設置することが望ましい。このマイクロホン11は、本発明の音声認識装置を実際に搭載するユーザ向け販売用の車種においてその設置位置がたとえばステアリング部分というように決められている場合にはその位置にマイクロホン11を設置して、そのマイクロホン11から雑音データを収集する。

【0040】

また、設計・開発段階などでマイクロホン11の設置位置が決定されていない場合は、設置候補となり得る複数の位置に設置して、それぞれのマイクロホン11から雑音データを収集するようにしてもよい。この実施の形態では、設置位置はすでに決められているものとし、その設置位置に設置された1つのマイクロホン11から雑音データを収集する例について説明する。

【0041】

また、マイクロホン11からの雑音収集と同時に、自動車の走行状況、現在位置、天候（ここでは雨量）、自動車に搭載されている各種機器の動作状況などを表す情報（雑音パラメータと呼ぶことにする）を時系列で収集する。

【0042】

この雑音パラメータは、自動車の速度、エンジン回転数、変速ギヤの位置、窓の開閉状況（開度）、エアコンの運転状況（風量の設定度合い）、ワイパーの動作状況、方向指示器の動作状況、雨量計からの雨量、GPS（Global Positioning System）による走行位置情報、カーオーディオの音信号などであり、これら各雑音パラメータは、これらの雑音パラメータの取得が可能な雑音パラメータ取得部13から時系列に取得されて雑音パラメータ記録部21に記録される。なお、これら雑音パラメータ取得部13は自動車に設置され、たとえば、走行速度を取得する速度情報取得部131、エンジン回転数を取得する回転数情報取得部132、変速ギヤ位置情報取得部133、窓の開度をたとえば開度0%、開度50%、開度100%などというような情報として取得する窓開度情報取得部134、エアコンの運転状態をたとえば運転停止・風量（弱・中・強）などという情報

として取得するエアコン運転情報取得部135、ワイパーのオン・オフ情報を取得するワイパー情報取得部136、方向指示器のオン・オフ情報を取得する方向指示器情報取得部137、GPSからの現在位置情報を取得する現在位置情報取得部138、雨量センサからの雨量情報（雨量なし、雨量少・雨量多など）を取得する雨量情報取得部139、カーオーディオからの音量などの音情報を取得するカーオーディオ情報取得部140などからなる。

【0043】

なお、上述したような実際に自動車を走行させてマイクロホン11から時系列に収集される雑音データと、雑音パラメータ取得部13の各情報取得部131～140から時系列に取得される各雑音パラメータは、実際に自動車を走行（停止状態も含む）させることによって得られるものである。

【0044】

すなわち、その自動車をたとえば1ヶ月あるいは数ヶ月といった長期間、色々な場所、色々な天候のもとで走行させ、かつ、各雑音パラメータを色々変化、たとえば、走行速度を色々変化させたり、回転数を色々変化させたり、変速ギヤを色々変化させたり、窓の開度を色々変化させたり、エアコンを色々な設定状態としたり、カーオーディオから色々な音信号を出力させたり、ワイパーや方向指示器などを適宜操作させるというように、自動車の走行時にあり得る様々な状態を作り出すことによって、マイクロホン11からは多種多様な雑音が時系列で入力され、入力信号処理部12で増幅処理やデジタル信号への変換処理（A/D変換）がなされて、雑音データとして雑音データ記録部22に記録されるとともに、その時々の各雑音パラメータが雑音パラメータ取得部13で時系列に取得されて雑音パラメータ記録部21に記録される。

【0045】

そして、クラスタリング処理部23が、マイクロホン11で収集された時系列の雑音データ（雑音データ記録部22に記録された時系列の雑音データ）を用いて、ある統計的手法によってその雑音データをクラスタリングしてn個の雑音クラスタを作成し、それぞれの雑音クラスタごとのn種類の雑音データ N_1 , N_2 , \dots , N_n を生成する。

【0046】

このクラスタリングには幾つかの手法が考えられるが、たとえば、収集された時系列の雑音データの特徴ベクトルをベクトル量子化して、そのベクトル量子化結果を用いて n 個の雑音クラスタにクラスタリングする方法、あらかじめ用意された幾つかの音声認識用データに実際に重畳させてそれを実際に音声認識させて、その認識結果に基づいて n 個の雑音クラスタにクラスタリングする方法などがある。

【0047】

なお、クラスタリングによって得られた n 種類の雑音データ $N1, N2, \dots, Nn$ は、それぞれの雑音データ $N1, N2, \dots, Nn$ ごとに、走行速度、回転数、変速ギヤ、窓の開度、エアコンの運転状況など上述した様々な雑音パラメータの値に依存したものであるため、これらそれぞれの雑音パラメータと n 種類の雑音データ $N1, N2, \dots, Nn$ は互いに対応付けられたものとなっている。

【0048】

たとえば、雑音データ $N1$ は、走行速度が時速 40 km から時速 80 km の範囲内、回転数が 1500 rpm から 3000 rpm の範囲内、変速ギヤがトップギヤ、窓の開度は 0 （閉状態）、エアコンは弱運転、ワイパーは OFF 、 \dots （そのほかの雑音パラメータは省略）のときに発生する雑音を表すデータであって、雑音データ $N2$ は、走行速度が時速 80 km から 100 km の範囲内、回転数が 3000 rpm から 4000 rpm の範囲内、変速ギヤがトップギヤ、窓の開度は 50% （半開き状態）、エアコンは強運転、ワイパーは OFF 、 \dots （そのほかの雑音パラメータは省略）のときに発生する雑音を表すデータであるといった具合である。

【0049】

これによって、現時点におけるそれぞれの雑音パラメータがそれぞれどのような値であるときは、そのときの雑音は n 種類の雑音データ $N1, N2, \dots, Nn$ のどの雑音データに属する雑音であるかを知ることができる。なお、このクラスタリングされた n 種類の雑音データ $N1, N2, \dots, Nn$ の具体例につ

いては後に説明する。

【0050】

このようにして、 n 種類の雑音データ $N_1 \sim N_n$ が作成されると、図1に示すように、これら雑音データ $N_1 \sim N_n$ を標準的な音声データ V (無響室などで収集されたたとえば、多数の話者が多種類単語について発話して得られた大量の音声データ) に重畳させ、 n 種類の雑音重畳音声データ VN_1, VN_2, \dots, VN_n を作成する。そして、この n 種類の雑音重畳音声データに対し、それぞれの雑音データ $N_1 \sim N_n$ を除去するに最適な雑音除去処理手法(前述したようにこの実施の形態では、3種類の雑音除去処理のどれか、またはそれらの組み合わせ)を用いて、雑音除去処理を行い、 n 個の雑音除去音声データ V_1', V_2', \dots, V_n' を作成し、この n 個の雑音除去音声データ V_1', V_2', \dots, V_n' を用いて音響モデルの学習を行い、 n 個の音響モデル M_1, M_2, \dots, M_n を作成する。

【0051】

この n 個の音響モデル M_1, M_2, \dots, M_n は、 n 種類の雑音データ $N_1 \sim N_n$ に対応するものである。つまり、音響モデル M_1 は雑音データ N_1 の重畳した音声データ(雑音重畳音声データ VN_1) から雑音データ N_1 を除去した後(雑音データ N_1 は完全には除去されないでその成分が残っている)の音声データ V_1' から作られた音響モデルであり、音響モデル M_2 は雑音データ N_2 の重畳した音声データから雑音データ N_2 を除去した後(雑音データ N_2 は完全には除去されないでその成分が残っている)の音声データから作られた音響モデルであり、音響モデル M_n は雑音データ N_n の重畳した音声データ(雑音重畳音声データ VN_n) から雑音データ N_n を除去した後(雑音データ N_n は完全には除去されないでその成分が残っている)の音声データ V_n' から作られた音響モデルである。

【0052】

以上のようにして本発明の実施の形態である自動車の機器操作を音声によって行う際の音声認識に用いられる音響モデル M_1, M_2, \dots, M_n が作成される。

【0053】

ところで、このような音響モデル M_1 , M_2 , ..., M_n を作成する際の雑音データ（マイクロホンから収拾された雑音データ）のクラスタリング処理について具体的に説明する。

【0054】

雑音データ収集を行うために長期間自動車を走行させることによって収集された雑音データには様々な雑音データが含まれる。たとえば、タイヤのパターンノイズ（主に速度に関係する）やエンジン音（主に速度、エンジンの回転数、ギヤ位置に関係する）、窓が開いている場合の風切り音、エアコンの運転音、雨が降っていれば雨そのものの音やワイパーの動作音、方向変換時には方向指示器の動作音、トンネル通過時は反響音、カーオーディオの動作中には音楽などの音信号などが雑音データとして収集される。

【0055】

そして、ある時刻においては、これらがすべて雑音データとして収集される場合もあり、また、ある時刻においては、これらのうち、たとえば、タイヤのパターンノイズやエンジン音だけが雑音データとして収集される場合もある。また、このような雑音とともに、それぞれの時刻対応にその自動車に設置された各種の雑音パラメータ取得部13で取得された雑音パラメータが記録されている。

【0056】

本来、雑音パラメータは前述したように多種存在し、マイクロホン11からは個々の雑音パラメータに対応した雑音、また、複数の雑音パラメータの組み合わせに対応した数多くの種類の雑音が収集され、このマイクロホン11から収集された雑音データを統計的手法によって実用的な数の雑音データとするためにクラスタリング処理して n 個の雑音クラスタとし、それぞれの雑音クラスタごとの雑音データ $N_1 \sim N_n$ を作成しているが、ここでは、説明を簡素化するために、雑音パラメータを3種類の雑音パラメータ（走行速度、エアコンの運転状態、雨量）のみで考え、これら走行速度、エアコンの運転状態、雨量の3つの雑音パラメータを3次元座標における3つの直交軸（ x 軸、 y 軸、 z 軸）上の値（ここではそれぞれ3段階の状態を示す値）で表してクラスタリングする例について説明す

る。

【0057】

すなわち、速度は「停車時（速度0）」、「低速時」、「高速時」の3段階で表し、エアコンの運転状態は、「停止」、「弱風」、「強風」の3段階で表し、雨量は「無」、「少量」、「多量」の3段階で表すものとする。なお、速度の「低速」および「高速」は、たとえば、60 km/hまでを低速、それ以上を高速というように予め範囲を決めておく。また、雨量も同様に、雨量計から得られる1時間雨量が5 mmまでを少量、それ以上を多量というように予め範囲を決めておく。

【0058】

図3は上述の3種類の雑音パラメータによる雑音を収集対象とし、これら3種類の雑音パラメータに対応して発生する雑音を1つのマイクロホン11を用いて長期間かけて収集した雑音データ（これをNで表す）を1つの大きな球体で表している。この図3は速度を「停車時」、「低速」、「高速」の3段階、エアコンの運転状態を「停止」、「弱風」、「強風」の3段階、雨量を「無」、「少量」、「多量」の3段階として、それらを3次元座標上で表したものである。

【0059】

この雑音データNに対しベクトル量子化などを用いた統計的手法を用いることなく単純に、個々の雑音パラメータごとにクラスタリングすると、図4のようになる。この場合、3の3乗個（27個）の雑音クラスタが得られ、それぞれの雑音クラスタに対応した27個の雑音データN1～N27が得られる。この27個の雑音データN1～N27をそれぞれ小さな球体で表している。

【0060】

この図4において、幾つかの雑音データについて説明すると、たとえば、雑音データN1は速度が「停車時（速度0）」、エアコンは「停止」、雨量は「無」に対応する雑音データであり、雑音データN5は速度が「低速」、エアコンは「弱」、雨量は「無」に対応する雑音データであり、雑音データN27は速度が「高速」、エアコンは「強」、雨量は「多量」に対応する雑音データである。

【0061】

なお、この図4では、個々の雑音データN1～N27を雨量が「無」、「少量」、「多量」で色の濃さを分けして表しており、雨量が「無」における3×3個の雑音データN1～N9を最も薄い色で表し、雨量が「少量」における3×3個の雑音データN10～N18を中程度の濃さで表し、雨量の「多量」における3×3個の雑音データN19～N27を最も濃い色で表している。

【0062】

この図4によれば、確かに、自動車の現時点における雑音パラメータによって、マイクロホンにはどのような種類の雑音データが入力されるかを知ることができ、それによって、最適な音響モデルを用いて音声認識することが可能となる。たとえば、現在の自動車の速度が「低速」でエアコンが「弱運転」で雨量「無」であれば、そのときの雑音データはN5であり、その雑音データN5に対応した音響モデルを用いて音声認識を行う。

【0063】

この図4の場合は、マイクロホン11から得られた時系列の雑音データを単純に個々の雑音パラメータのとり得る状況の数（この例では27種類）ごとにクラスタリングした例であるが、これをある統計的手法によってクラスタリングした例について図5により説明する。

【0064】

このようなある統計的手法を用いてクラスタリングする例として、前述したように、雑音データの各時刻対応の特徴ベクトルをベクトル量子化して、そのベクトル量子化結果を用いて、n個の雑音クラスタにクラスタリングする方法、あらかじめ用意された幾つかの音声認識用データに実際に重畳させてそれを実際に音声認識させて、その認識結果に基づいてn個の雑音クラスタにクラスタリングする方法などがある。

【0065】

このような手法によってクラスタリングされた結果、図5に示すように9個の雑音クラスタが作成され、それぞれの雑音クラスタに対応した9種類の雑音データN1～N9が作成されたとする。

【0066】

この図5の場合、雨の音（雨量）が音声認識する上での雑音データとして最も影響が大きく、その次に自動車の走行速度の影響が大きく、エアコンの影響は雨や走行速度に比べて影響が小さいことを示している。

【0067】

この図5では雨が降っていなければ（雨量「無」）、自動車の走行速度が0（「停車時」）においては、エアコンの運転状態に対応した雑音データN1、N2、N3が作成される。また、自動車の走行速度が「低速」においては、エアコンの運転状態が「停止」に対応した雑音データN4とエアコンの運転状態が「弱」・「強」で1つの雑音データN5が作成される。すなわち、自動車がある速度で走行している場合には、エアコンの運転状態が「弱」であっても「強」であってもその運転音は、自動車の走行による雑音に比べるとその影響は殆どないと判断された結果である。また、自動車の速度が「高速」においては、エアコンの運転状態に関係なく、1つの雑音データN6が作成される。

【0068】

また、雨が降っていると、たとえ、雨量が「少量」であっても、エアコンがどのような運転状態にあるかは関係ないと判断され、自動車の走行速度に依存した雑音データが作成される。すなわち、雨量が「少量」である場合には、走行速度が「低速」まで（停車時を含む）に対応した雑音データN7と、「高速」に対応した雑音クラスタN8の2種類の雑音クラスタが作成されている。

【0069】

また、雨が「多量」である場合には、エアコンの運転状態や自動車の走行速度の影響は殆どないと判断され、1つの雑音クラスタN9が作成されている。

【0070】

以上のように、3種類の雑音パラメータ（走行速度、エアコンの運転状態、雨量）に対応する雑音を収集対象とし、これら3種類の雑音パラメータに依存する雑音を1つのマイクロホン11を用いて長期間かけて収集した雑音データNを、ある統計的手法でクラスタリングした結果、図5のような雑音クラスタN1～N9が作成されたとする。

【0071】

なお、この図5により得られた雑音クラスタは、説明をわかり易くするため、雑音パラメータを3つ（速度、エアコンの運転状態、雨量）とした例であるが、実際には、雑音パラメータは、前述したように、多種類存在し、これら多種類の雑音パラメータに依存した各種雑音を長期間収集して、時系列の雑音データを得て、その時系列の雑音データを統計的手法によってクラスタリングし、 n 個の雑音クラスタを作成して、これらそれぞれの雑音クラスタに対応する n 種類の雑音データ $N_1 \sim N_n$ を作成する。ここで、実用的な雑音クラスタ数としては、音響モデル作成処理や音声認識処理の効率化などの点から数個から10数個程度が好ましいが、これは、任意に設定することができる。

【0072】

このようにして、 n 個の雑音クラスタに対応する n 種類の雑音データ $N_1 \sim N_n$ が作成されると、前述したように（図1参照）、この n 種類の雑音データ $N_1 \sim N_n$ を標準的な音声データに重畳させて n 個の雑音重畳音声データ $V_{N1}, V_{N2}, \dots, V_{Nn}$ を作成し、この n 種類の雑音重畳音声データ $V_{N1}, V_{N2}, \dots, V_{Nn}$ に対し、それぞれの雑音を除去するに最適な雑音除去処理手法を用いて、雑音除去処理を行い、 n 種類の雑音除去音声データ V_1', V_2', \dots, V_n' を作成し、この n 種類の雑音除去音声データ V_1', V_2', \dots, V_n' を用いて音響モデルの学習を行い、 n 種類の音響モデル M_1, M_2, \dots, M_n を作成する。これによって、 n 種類の雑音データ N_1, N_2, \dots, N_n に対応した n 種類の音響モデル M_1, M_2, \dots, M_n を作成することができる。

【0073】

次に上述のようにして作成された n 種類の音響モデル M_1, M_2, \dots, M_n を用いた音声認識について説明する。

【0074】

図6は本発明の音声認識装置の構成図であり、機器操作用の音声コマンドや様々な雑音を入力する音入力手段としてのマイクロホン11、このマイクロホン11から入力された音声コマンドを増幅するとともにデジタル信号に変換（A/D変換）する入力信号処理部12、自動車に搭載されている各種機器群の動作状

況を雑音パラメータとして取得する雑音パラメータ取得部13、この雑音パラメータ取得部13から取得された各種の雑音パラメータに基づいて、現在の雑音の種類が前述したn個にクラスタリングされたn種類の雑音データN1～Nnのどれに相当するかを判別する雑音種類判別部14、雑音データN1～Nnそれぞれに最適な雑音除去方法が保存された雑音除去方法保存部15、雑音種類判別部14で判別された雑音データに最適な雑音除去方法を雑音除去方法保存部15に保存された各種の雑音除去方法の中から選択し、マイクロホン11から入力された音声データ（デジタル変換後の雑音重畳音声データ）に対して雑音除去処理を行う雑音除去処理部16、この雑音除去処理部16によって雑音除去された雑音除去音声データに対し、前述の手法によって作成された音響モデルM1～Mn（n種類の雑音データN1～Nnに対応）を用いるとともに、言語モデル17を用いて音声認識を行う音声認識処理部8を有した構成となっている。

【0075】

なお、雑音種類判別部14は、雑音パラメータ取得部13からの情報を受け取って、現在、どのような雑音が発生しているかの判別を行い、それによって、現在の雑音状況が雑音データN1～N9のどの雑音データに属するかを判定する。

【0076】

すなわち、雑音パラメータ取得部13は、前述したように、たとえば、走行速度を取得する速度情報取得部131、エンジン回転数を取得する回転数情報取得部132、変速ギヤ位置情報取得部133、窓の開度をたとえば開度0%、開度50%、開度100%などというような情報として取得する窓開度情報取得部134、エアコンの運転状態をたとえば運転停止・風量（弱・中・強）情報として取得するエアコン運転情報取得部135、ワイパーのオン・オフ情報を取得するワイパー情報取得部136、方向指示器のオン・オフ情報を取得する方向指示器情報取得部137、GPSからの現在位置情報を取得する現在位置情報取得部138、雨量センサからの雨量情報（雨量なし、雨量少・雨量多など）を取得する雨量情報取得部139、カーオーディオからの音量などの音情報を取得するカーオーディオ情報取得部140などからなるので、雑音パラメータ取得部13は、これら各情報取得部131～140からの情報に基づいて、現在、どのような雑

音が発生しているかの判別を行い、それによって、現在の雑音状況が $N_1 \sim N_n$ の雑音データのどの雑音データに属するかを判定する。

【0077】

たとえば、現在の走行速度が時速60km、エンジン回転数が2000rpm、ギヤの位置がトップギヤ、窓の開度が0%（閉状態）、ワイパーはオフ、方向指示器はオフといった情報を雑音パラメータとして雑音種類判別部14が受け取ると、これらの雑音パラメータから、現在の雑音種類を判別し、それによって、現在の雑音が雑音データ $N_1 \sim N_n$ のどの雑音データに属するかを判定し、仮に、現在の雑音が雑音データ N_1 に属すると判定されたとすると、その判定結果を雑音除去処理部16と音声認識処理部18に送る。

【0078】

雑音除去処理部16は、雑音種類判別部14からの現在の雑音種類を示す情報を受け取ると、入力信号処理部12からの雑音重畳音声データに対して最適な雑音除去方法を用いた雑音除去処理を行う。たとえば、雑音種類判別部14からの現在の雑音種類を示す情報が雑音データ N_1 に属する雑音であることを示す情報であるとすれば、雑音データ N_1 に最適な雑音除去手法が雑音除去方法保存部15から選択され、その雑音除去方法によって雑音重畳音声データに対する雑音除去処理がなされる。

【0079】

なお、この雑音除去処理は、この実施の形態の場合、前述したように、たとえば、スペクトラム・サブトラクション（SS）やケプストラム平均化処理（CMN）などのいずれかあるいはそれらの組み合わせのどれかによって行う。また、現在の雑音の中にカーオーディオからの音信号やワイパーの動作音、方向指示器の動作音が含まれる場合は、これらの雑音を直接除去するような処理も可能とする。

【0080】

たとえば、マイクロホン11に入力された雑音重畳音声データに含まれるカーオーディオからの音信号に対しては、カーオーディオから直接得られる音信号、すなわち、カーオーディオ情報取得部140から得られるカーオーディオ信号を

雑音除去処理部 16 に与え（図 6 において一点鎖線で示す）、そのカーオーディオ信号をマイクロホンに入力された雑音重畳音声データから差し引くことで、マイクロホン 11 に入力された雑音重畳データに含まれるカーオーディオの音信号成分を除去することができる。このとき、雑音除去処理部 16 では、マイクロホン 11 からの雑音重畳音声データに含まれるカーオーディオ信号は、カーオーディオから直接得られる信号に比べて一定の時間遅れが生じるので、その時間遅れ分を考慮した除去処理を行う。

【0081】

また、ワイパーや方向指示器の動作音は、周期的な動作音であり、それぞれの周期や雑音成分（動作音）は車種によって決められているので、その周期に対応したタイミング信号（図 6 において一点鎖線で示す）をワイパー情報取得部 136 や方向指示器情報取得部 137 から雑音除去処理部 16 に送ることで、雑音除去処理部 16 では、そのタイミングでワイパーの動作音や方向指示器の動作音を除去することができる。この場合も、マイクロホン 11 からの雑音重畳音声データに含まれるワイパーの動作音や方向指示器の動作音は、ワイパーや方向指示器から直接得られる動作信号に比べて一定の時間遅れが生じるので、その時間遅れ分を考慮したタイミングで雑音除去処理を行う。

【0082】

以上のようにして、マイクロホン 11 から入力されたある時刻の雑音重畳音声データ（音声コマンドとその時点でマイクロホンに入力される雑音からなる）に対する雑音除去処理がなされると、その雑音の除去された雑音除去音声データは音声認識処理部 18 に送られる。

【0083】

音声認識処理部 18 には、雑音種類判別部 14 から雑音データの種類を示す情報が与えられており、その雑音データの種類に対応する音響モデルを選択して用いるとともに、言語モデル 17 を用いて音声認識処理を行う。

【0084】

たとえば、マイクロホン 11 に入力された話者からの音声コマンド（たとえば、自動車内の何らかの機器を音声によって操作するための音声コマンド）に重畳

された雑音データは雑音クラスタN1に属する雑音であると判定されたとすれば、音声認識処理部18では、音響モデルとしては、雑音クラスタN1に対応する音響モデルM1を用いる。

【0085】

この音響モデルM1は、前述の音響モデル作成方法で説明したように、雑音クラスタに属する雑音データN1を音声データに重畳させ、その雑音重畳音声データから雑音除去して雑音除去音声データを作成し、その雑音除去音声データから作成された音響モデルであるため、話者の発した音声コマンドに重畳された雑音が雑音データN1に属する場合には、その音声コマンドにとって最適な音響モデルとなり、認識性能を高めることができる。

【0086】

1つの具体例として、図5に示すような9個の雑音クラスタに対応する9種類の雑音データN1～N9が作成され、これら9種類の雑音データN1～N9に対応した音響モデルM1～M9が作成されている場合の音声認識動作について説明する。この場合、雑音パラメータは、速度、エアコンの運転状態、雨量のみについて考え、これらの雑音パラメータを長期間かけて収集することによって、図5に示すような9個の雑音クラスタに対応する9種類の雑音データN1～N9が作成され、それに対応した音響モデルM1～M9が作成されたとする。

【0087】

今、運転者が運転中に音声コマンドを与えると、音声認識装置側ではその音声コマンドを認識して、その認識結果に基づいた機器操作を行う場合を考える。なお、このときの3種類の雑音パラメータ（速度、エアコンの運転状態、雨量）は、走行速度が時速40km（低速走行であるとする）、エアコンの運転状態は「弱」、雨量は「無」とする。

【0088】

この場合、自動車内のある位置（ステアリングなど）に設置されたマイクロホン11には、そのときの状況に応じた雑音が入力され、その状態で運転者が何らかの音声コマンドを発すると、その音声コマンドにはそのときの状況に応じた雑音データが重畳され、その雑音重畳音声データは入力信号処理部12で増幅やA

／D変換されたのちに雑音除去処理部12に送られる。

【0089】

一方、雑音種類判別部14には、この場合、現在の雑音パラメータとして、雑音パラメータ取得部13の速度情報取得部131から現時点の走行速度情報、エアコン運転情報取得部135からエアコンの運転情報、雨量情報取得部139からの雨量情報が雑音パラメータとして与えられており、それらの雑音パラメータに基づいて、現在の雑音状況を判断して現在の雑音がどの雑音クラスタの雑音データであるかを判定する。

【0090】

この場合、速度計から走行速度は時速40km、エアコンの運転状態は「弱」、雨量計からの雨量は「無」であるので、図5から現在の雑音は雑音データN5であると判定され、それを示す情報を雑音除去処理部16と音声認識処理部18に送る。

【0091】

これによって、雑音除去処理部16では、その雑音データN5に最適な雑音除去処理手法を用いて雑音除去処理を行い、その雑音除去音声データを音声認識処理部18に送る。

【0092】

音声認識処理部18では、雑音種類判別部14から送られてきた雑音データN5に対応した音響モデルM5を選択し、その音響モデルM5と言語モデル17を用いて、雑音除去処理部16で雑音除去された雑音除去音声データに対して音声認識処理を行う。そして、この音声認識結果に基づいて機器操作を行う。この機器操作の例としては、たとえば、ナビゲーションシステムに対する目的地の設定などである。

【0093】

以上のように、この実施の形態における音声認識装置では、音声コマンドに重畳された雑音の種類（雑音データN1～N5）に対応した雑音除去処理手法（音響モデル作成時と同じ雑音除去処理手法）を用いて雑音除去を行い、その雑音除去のなされた音声データ（雑音除去音声データ）に対し、最適な音響モデルを用

いて音声認識を行うようにしている。すなわち、自動車のその時々¹の走行状況、走行位置、車載機器の動作状況などに対応した様々な種類の雑音が音声コマンドに重畳されても、それに対応した最適な雑音除去が行え、その雑音除去後の音声データに対し、最適な音響モデルを用いて音声認識することができるので、様々な雑音環境下において高い認識性能を得ることができる。

【0094】

特に、自動車などの場合、車種が限定されているような場合に特に有効なものとなる。すなわち、雑音データ収集を行って音響モデルを作成するための雑音データ収集用車種と実際に本発明の音声認識装置を搭載するユーザ向け販売用車種が同じであれば、雑音データ収集用車種における雑音データ収集のためのマイクロホン取り付け位置とユーザ向け販売用車種における音声コマンド入力用のマイクロホン取り付け位置を同じとすることで、雑音データがほぼ同じ条件でマイクロホンから入力されるので、適正な音響モデルを選択することができ、高い認識性能を得ることができる。

【0095】

なお、音響モデルを作成するための雑音データ収集用自動車は専用に用意することも可能であるが、ユーザ向け販売用自動車に音響モデル作成（図2で示した雑音データN1～Nn作成を含む）を行うための機能を図6の音声認識機能とともに搭載して、1台の自動車で音響モデル作成機能と音声認識機能の両方を可能とすることも可能である。その場合、マイクロホン11、入力信号処理部12、雑音パラメータ取得部13、雑音除去処理部15などは音響モデル作成時と音声認識時で共用することができる。

【0096】

このように、ユーザ向け販売用自動車に音響モデル作成機能と音声認識機能の両方を持たせることで、雑音環境の変化などにより雑音のクラスタリングの変更が容易に行え、それに伴って、音響モデルを新たに生成したり、更新したりすることが可能となり、雑音環境の変化への対応がし易くなる。

【0097】

なお、本発明は上述の実施の形態に限られるものではなく、本発明の要旨を逸

脱しない範囲で種々変形実施可能となるものである。たとえば、前述の音声認識装置の実施の形態において、雑音種類判別処理部は、自動車の現時点における雑音パラメータを入力することによって現在の雑音がどの n 種類の雑音クラスタのうちのどの雑音クラスタに属するかを判定するようにしているが、この雑音種類判定を行う際、図 7 に示すように、雑音修理判定部には雑音パラメータのほかに、音データの重畳された雑音重畳音声データ（デジタル変換後の雑音重畳音声データ）を入力して、この雑音重畳音声データと各種雑音パラメータとによって、現在の雑音がどの雑音クラスタに属する雑音であるかを判定するようにしてもよい。

【0098】

このように、マイクロホン 11 から入力される雑音重畳音声データを雑音種類判別処理部 14 に入力させることによって、現在の S/N 比をより正確に判別し易くなり、各音響モデル $NM1 \sim Mn$ が S/N 比の大きさも考慮した音響モデルが作成されている場合、現在の S/N 比に応じた最適な音響モデルを選択することができ、より、適正な音声認識を行うことができる。

【0099】

また、雑音パラメータの種類は、前述の実施の形態で説明した種類に限られるものではなく、そのほかにも種々用いることができる。なお、音響モデル作成を行うために、実際に自動車を長期間走行させて、各種の雑音パラメータの収集を試みて、統計的手法によりクラスタリングすることによって複数個の雑音データを作成したとき、ある雑音パラメータがクラスタリングに影響を与えるものでないと判断される場合もあるが、その場合は、音声認識時において、その雑音パラメータは、雑音種類判定部が雑音種類を判定する際の雑音パラメータから除外することができる。

【0100】

また、前述の実施の形態では、自動車を例にとって説明したが、自動車に限られるものではなく、運転者の操作によって自走可能な車両（たとえば、オートバイなどの二輪車など）にも適用できることは言うまでもない。

【0101】

また、本発明は以上説明した本発明を実現するための処理手順が記述された処理プログラムを作成し、その処理プログラムをフロッピーディスク、光ディスク、ハードディスクなどの記録媒体に記録しておくこともでき、本発明は、その処理プログラムの記録された記録媒体をも含むものである。また、ネットワークから当該処理プログラムを得るようにしてもよい。

【0102】

【発明の効果】

以上説明したように、本発明の音声認識用の音響モデル作成方法によれば、自動車などの車両をある期間走行させることによって、その車両内で収集可能な各種雑音を時系列で収集し、その収集された雑音をある統計的手法によってクラスタリングしてn種類の雑音データを作成し、そのn種類の雑音データを予め用意された標準的な音声データに重畳させてn種類の雑音重畳音声データを作成し、そのn種類の雑音重畳音声データに対して雑音除去処理を行ってn種類の雑音除去音声データを作成し、そのn種類の雑音除去音声データからn種類の音響モデルを作成するようにしているので、車両の走行状況、周囲の環境、天候、車両内の機器の動作状況に起因する様々な雑音の種類に対応した最適な音響モデルを作成することができる。

【0103】

また、本発明の音声認識装置によれば、現在の雑音が予めクラスタリングされたn種類の雑音データのうちの雑音データに属する雑音であるかを判別する雑音種類判別し、また、音声入力手段から雑音の重畳された雑音重畳音声を入力するとともに、前記雑音種類判定結果を入力し、雑音重畳音声に対してその雑音種類判定結果に基づいて雑音除去を行う雑音除去を行い、雑音除去された雑音除去音声に対し、前記n種類の音響モデルのうち前記雑音種類判定部で判定された雑音種類に対応した音響モデルを用いて音声認識を行うようにし、かつ、上述した音響モデル作成方法および音響モデル作成処理プログラムによって作成されたn種類の音響モデルを持っているので、自動車など車両の走行状況、走行場所による周囲の環境、天候、車両に搭載されている機器の動作状況に起因する雑音に対し最適な雑音除去処理が行えとともに、その時点の雑音に対して最適な音響モ

デルを用いた音声認識が可能となるので、自動車の車内など特有の雑音環境下で高い認識性能を得ることができる。

【図面の簡単な説明】

【図 1】 本発明の音声認識用の音響モデル作成方法の大まかな処理手順を説明する図である。

【図 2】 収集された雑音データをクラスタリングしてそれぞれの雑音クラスごとの雑音データを生成する処理を説明する図である。

【図 3】 ある 3 種類の雑音パラメータに対応して発生する雑音を 1 つのマイクロホンを用いて長期間かけて収集した雑音データ N を 3 次元座標上で 1 つのデータで表した図である。

【図 4】 図 3 の雑音データ N に対し、単純に個々の雑音パラメータごとにクラスタリングしてそれぞれのクラスごとに作成された雑音データを示す図である。

【図 5】 雑音データ N に対しベクトル量子化などの統計的手法によってクラスタリングして得られたそれぞれのクラス対応に作成された雑音データを示す図である。

【図 6】 本発明の音声認識装置の構成図である。

【図 7】 図 5 で示した音声認識装置の変形例を説明する図である。

【図 8】 従来の音響モデル作成を概略的に説明する図である。

【図 9】 図 8 で作成された音響モデルを用いた従来の音声認識装置の概略的な構成図である。

【符号の説明】

- 1 雑音重畳音声データ作成部
- 2 雑音除去処理部
- 3 音響モデル学習処理部
- 1 1 マイクロホン
- 1 2 入力信号処理部
- 1 3 雑音パラメータ取得部
- 1 4 雑音種類判別部

15 雑音除去方法保存部

16 雑音除去処理部

18 音声認識処理部

21 雑音パラメータ記録部

22 雑音データ記録部

23 クラスタリング処理部

N_1, N_2, \dots, N_n 各雑音クラスタ対応の雑音データ

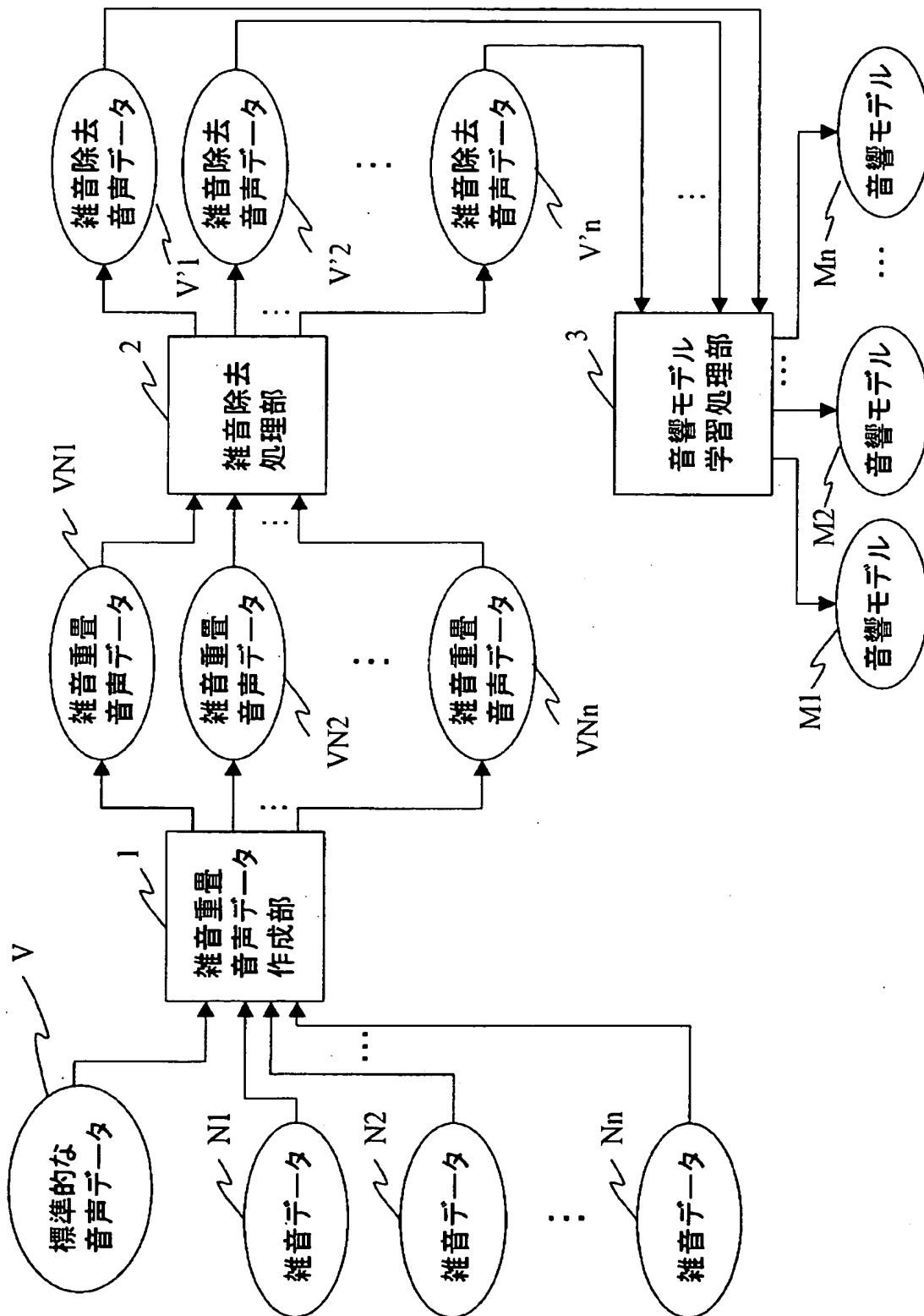
VN_1, VN_2, \dots, VN_n 雑音重畳音声データ

V_1', V_2', \dots, V_n' 雑音除去音声データ

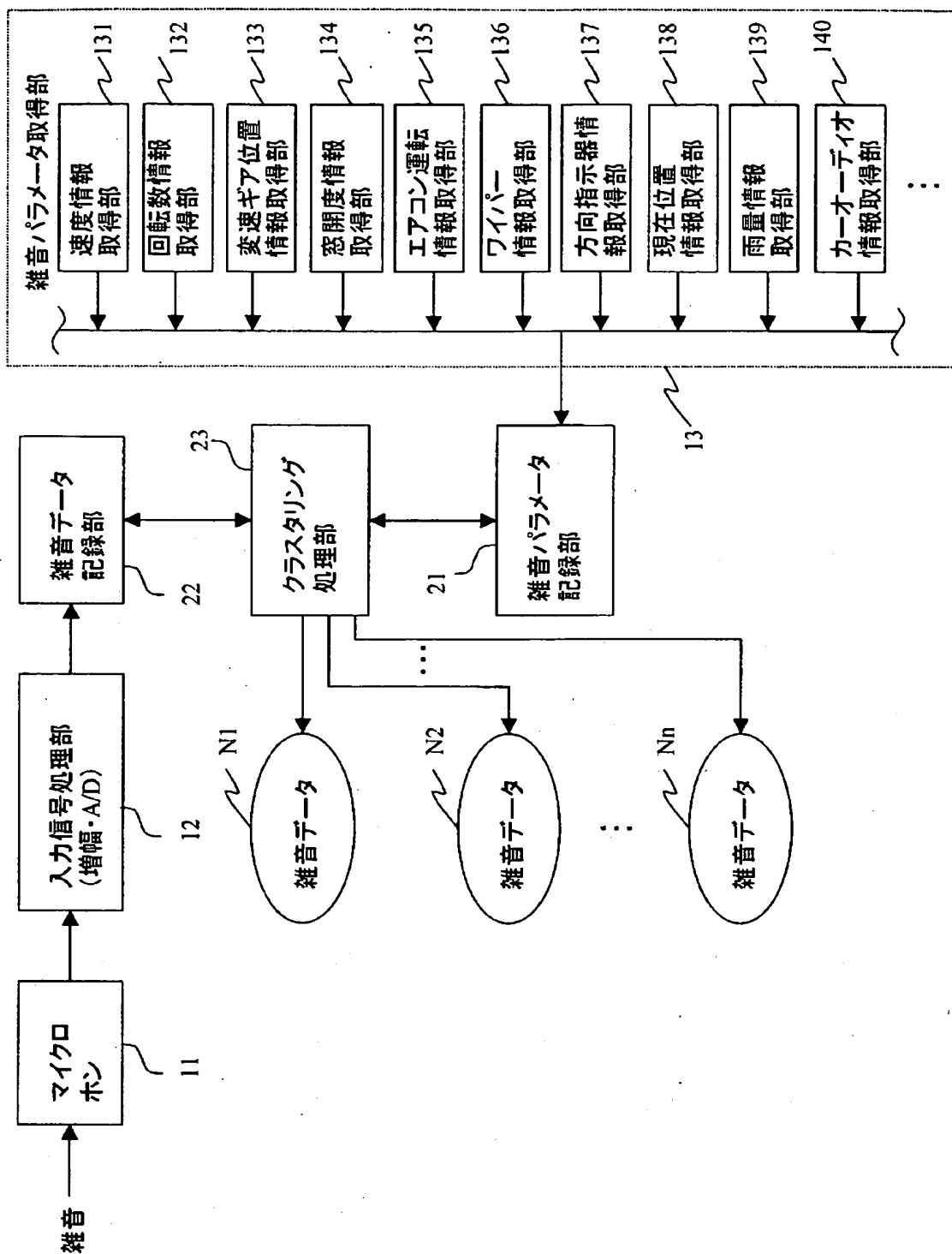
M_1, M_2, \dots, M_n 音響モデル

【書類名】 図面

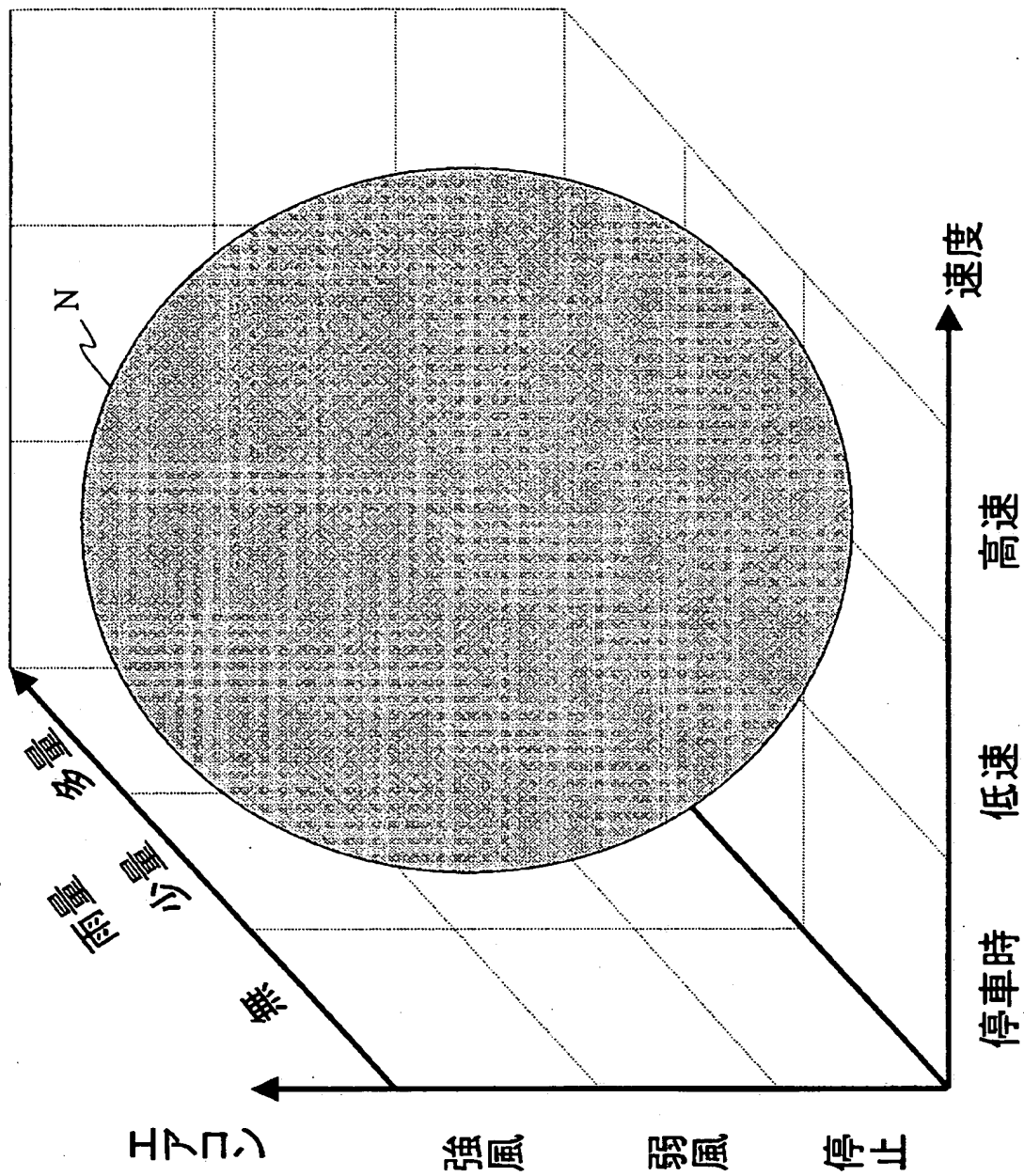
【図 1】



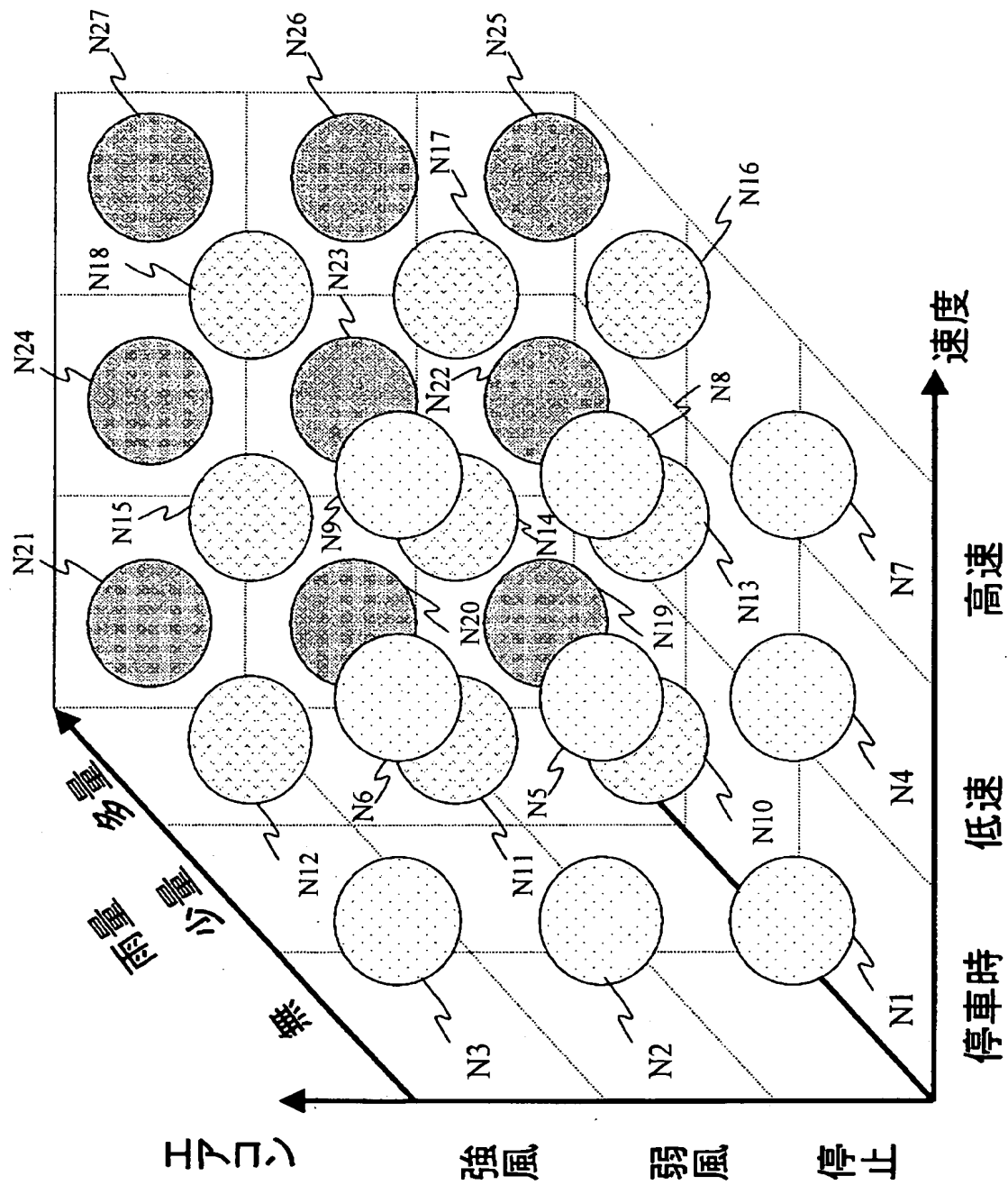
【図 2】



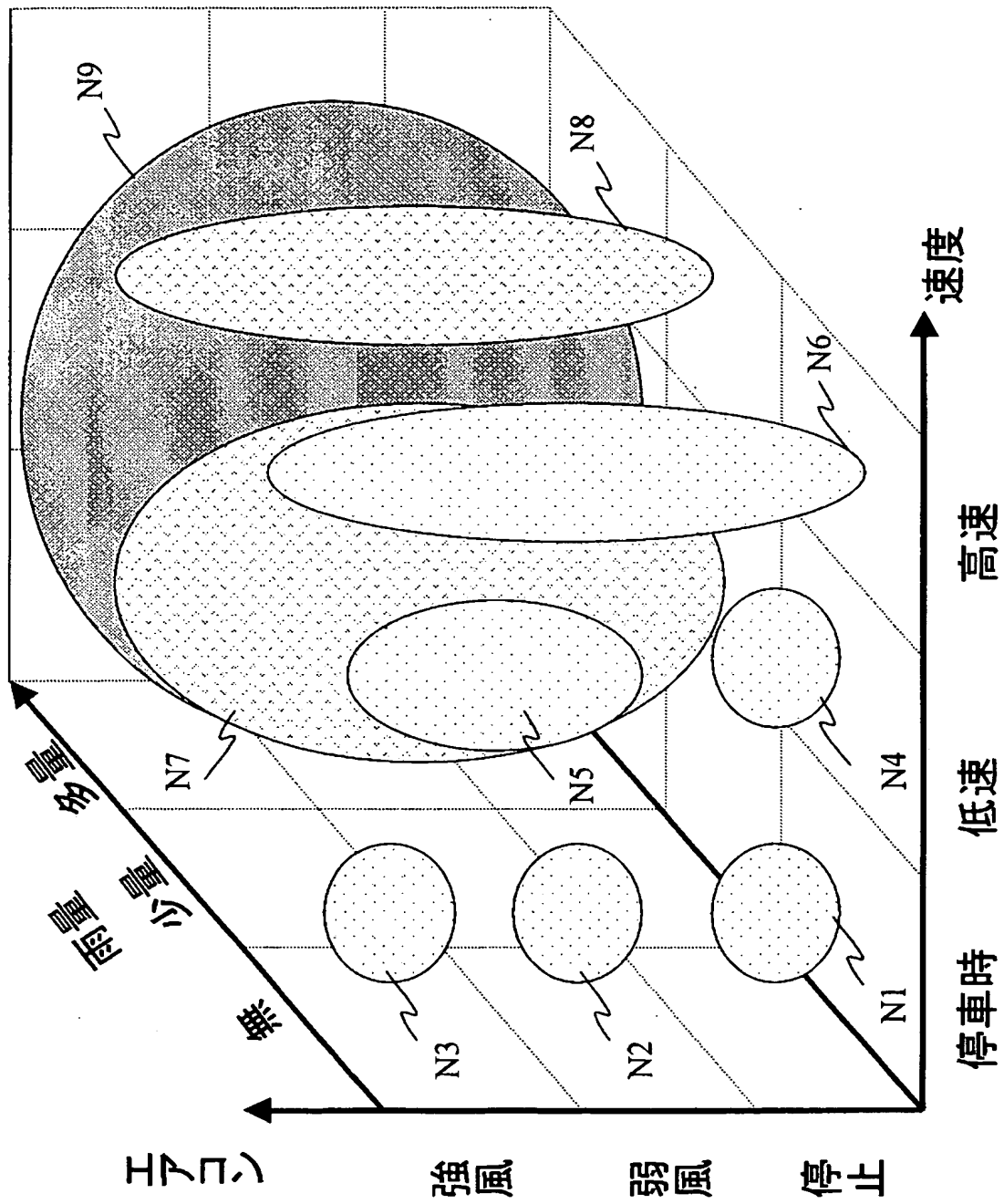
【図 3】



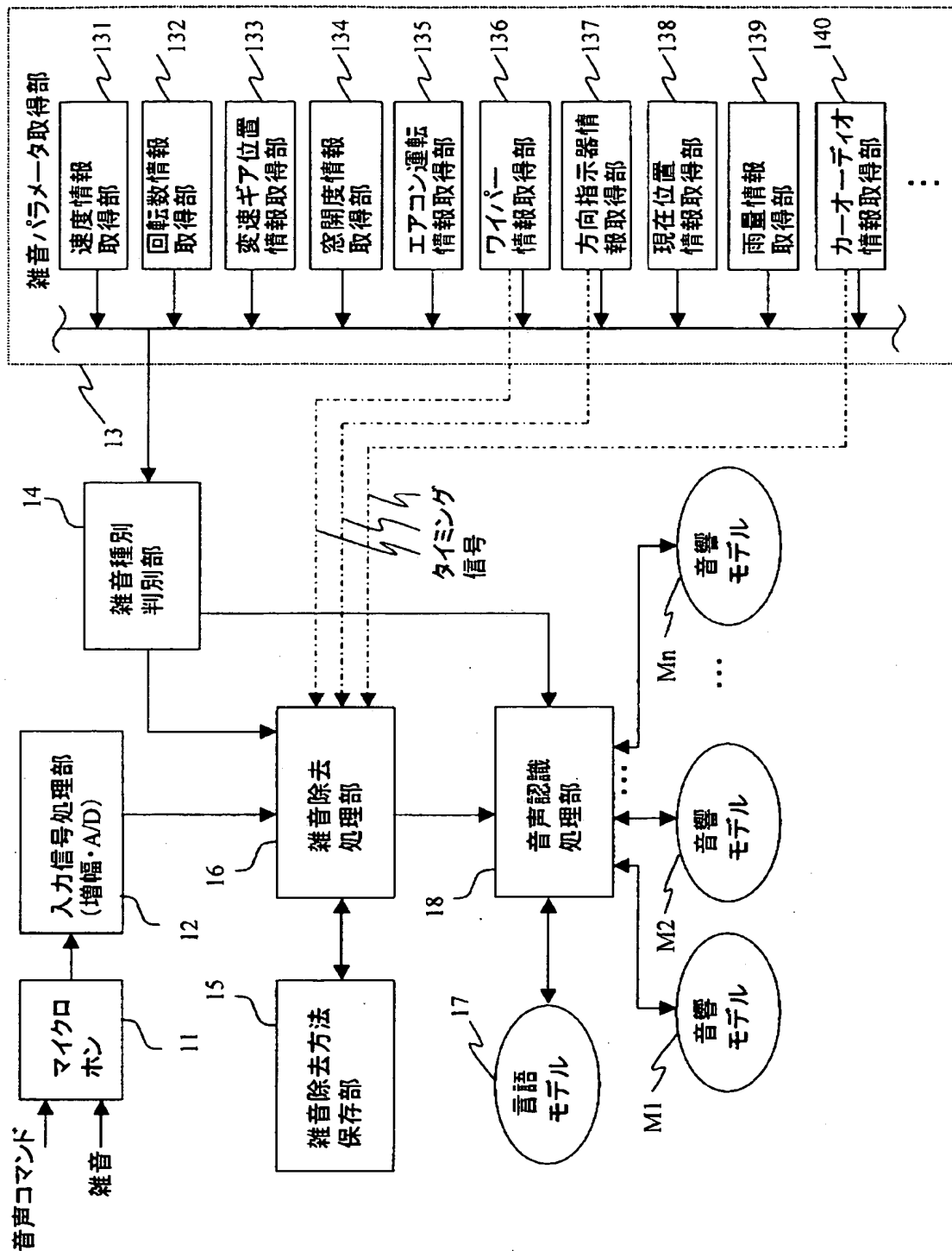
【図 4】



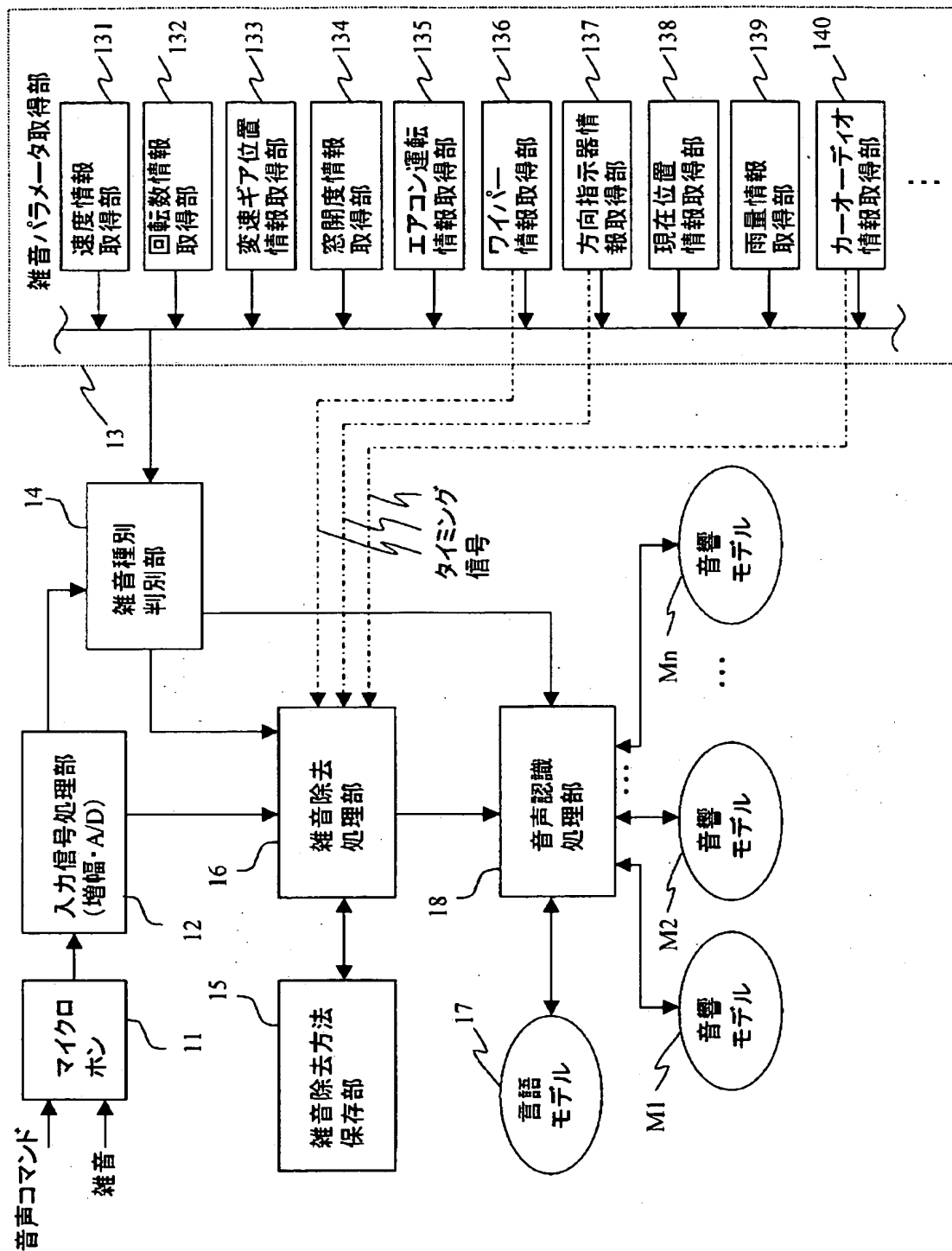
【図 5】



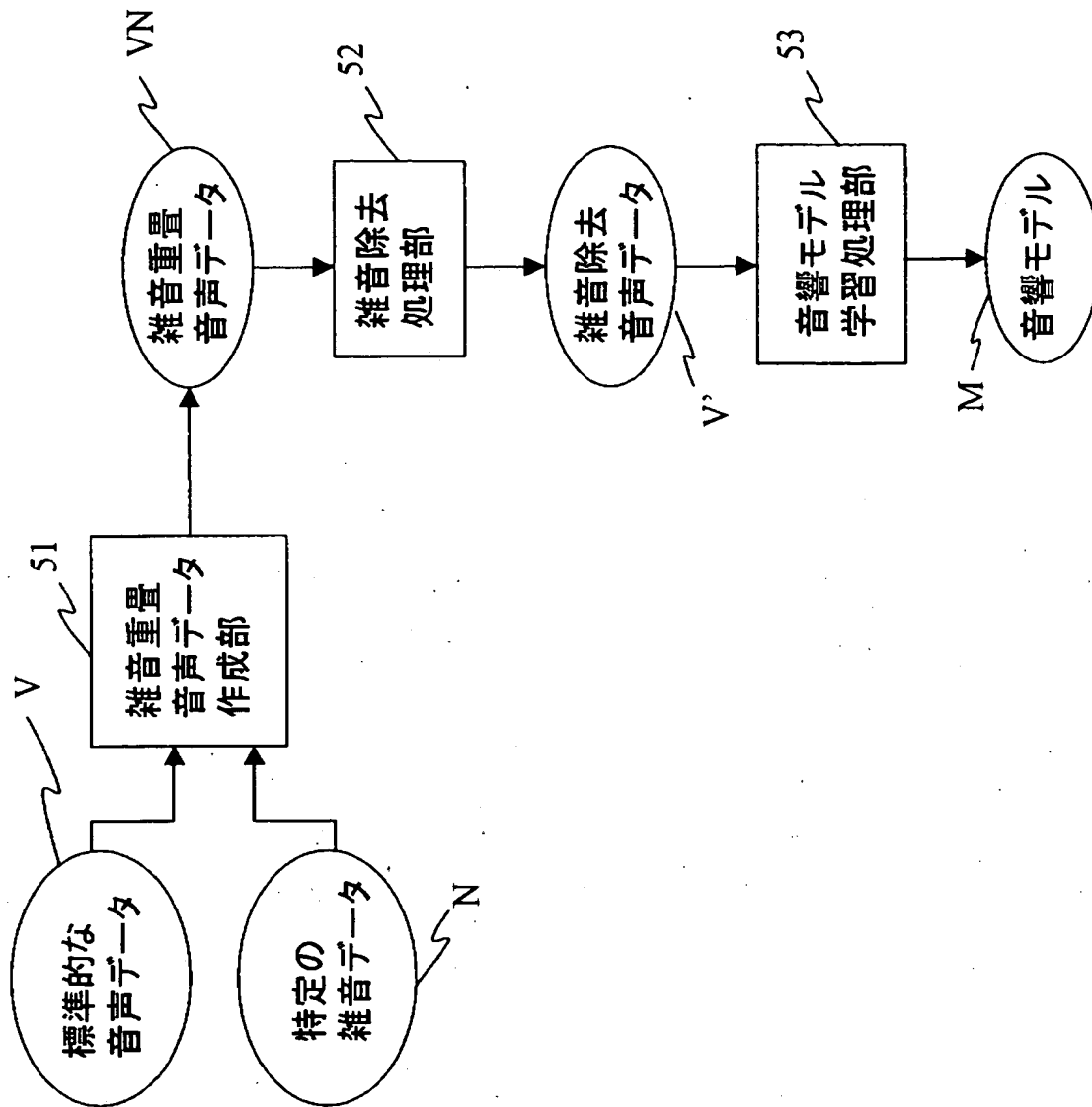
【図 6】



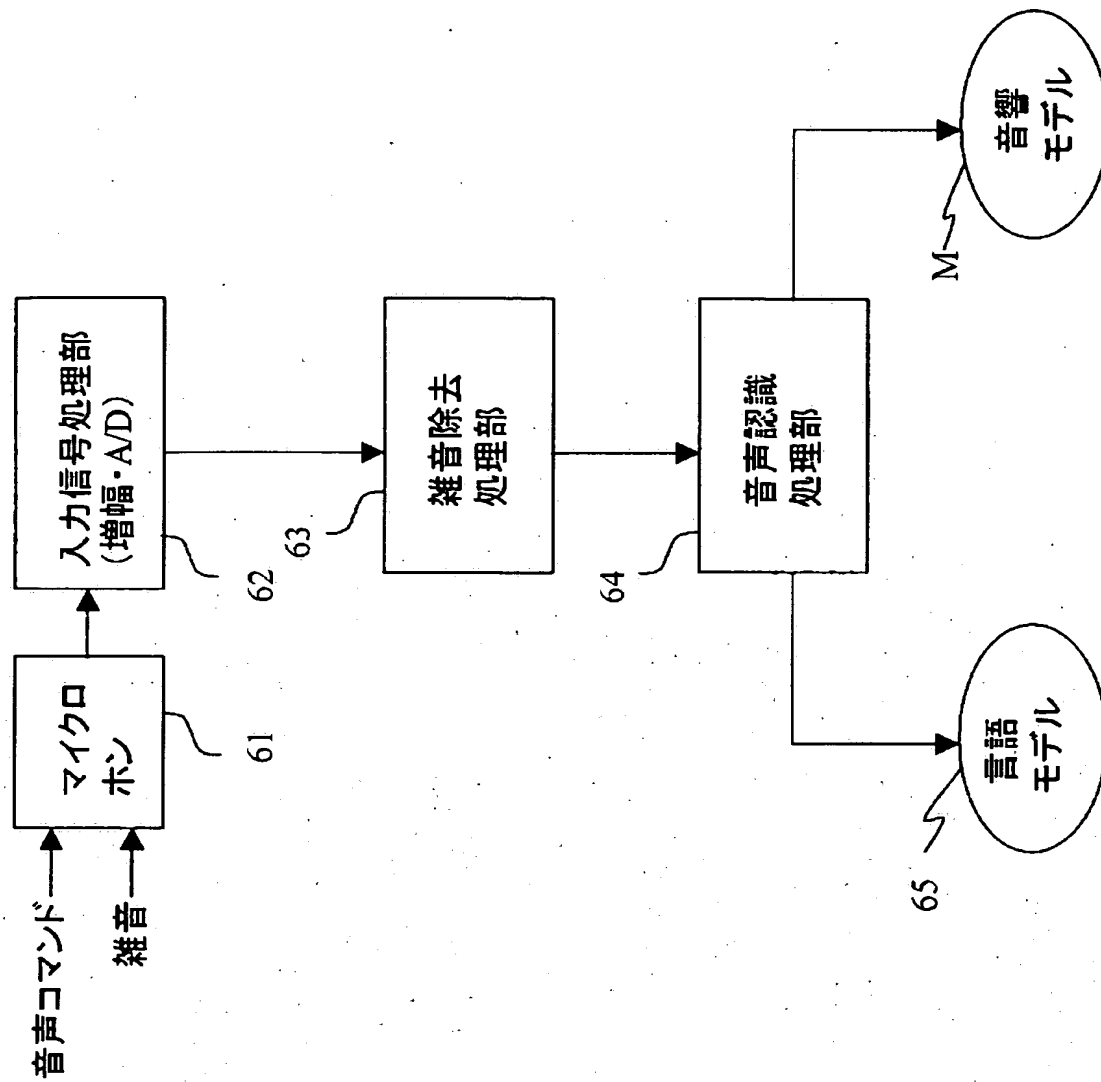
【圖 7】



【図8】



【図 9】



【書類名】 要約書

【要約】

【課題】 自動車内など多種多様な雑音環境下で高い認識性能を得る。

【解決手段】 自動車の走行状況、周囲の環境、搭載機器の動作状況を表すデータを入力し、それによって、現在の雑音が予めクラスタリングされた n 種類の雑音データのどの雑音データに属するかを判別する雑音種類判別部 14 と、 n 種類の雑音データをそれぞれ標準的な音声データに重畳させて n 種類の雑音重畳音声データを作成したのち、それぞれ雑音除去処理された n 種類の雑音除去音声データに基づいて作成された n 種類の音響モデル $M_1 \sim M_n$ と、マイクロホン 11 から雑音の重畳された雑音重畳音声を入力するとともに、前記雑音種類判定結果を入力し、雑音重畳音声に対して雑音除去を行う雑音除去処理部 16 と、その雑音除去音声に対し、 n 種類の音響モデルのうち雑音種類判定部 14 で判定された雑音種類に対応した音響モデルを用いて音声認識を行う音声認識処理部 18 を有する。

【選択図】 図 6

認定・付加情報

特許出願の番号	特願 2002-318627
受付番号	50201652766
書類名	特許願
担当官	第八担当上席 0097
作成日	平成14年11月 1日

<認定情報・付加情報>

【提出日】	平成14年10月31日
-------	-------------

次頁無

特願 2002-318627

出 願 人 履 歴 情 報

識別番号

[000002369]

1. 変更年月日

1990年 8月20日

[変更理由]

新規登録

住 所

東京都新宿区西新宿2丁目4番1号

氏 名

セイコーエプソン株式会社